

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-160292

(43)Date of publication of application : 23.06.1995

(51)Int.Cl.

G10L 7/04  
G10L 9/18

(21)Application number : 05-306898

(71)Applicant : SONY CORP

(22)Date of filing : 07.12.1993

(72)Inventor : AKAGIRI KENZO

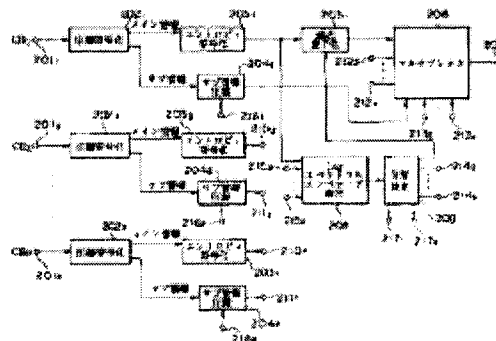
## (54) MULTILAYERED CODING DEVICE

### (57)Abstract:

**PURPOSE:** To improve compressibility and to perform bit distribution among channels by detecting energy for each channel and performing encoding in which information quantity of each channel is allocated with a ratio in accordance with this detected output.

**CONSTITUTION:** This device has compression-encoding circuits 201-202n and entropy encoding circuit 2031-203n being a first encoding means in which digital audio signals of each channel CH1-CHn are inputted and the digital audio signals are compression-encoded. This device has also a second encoding means in which an output encoded signal of the first encoding means is supplied, energy for each channel is detected, and encoding is performed in which

information quantity of each channel is allocated. This second encoding means has a Log spectrum envelop detecting circuit 208, a distribution deciding circuit 209, an adaptive quantization circuits 2031-203n, and a sub-information compression circuit 2041-204n. Thereby, bit distribution among channels and high compressibility can be realized.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-160292

(43) 公開日 平成7年(1995)6月23日

(51) Int. Cl. <sup>6</sup>	識別記号	片内整理番号	P I	技術表示箇所
G 1 0 L 7/04	G			
9/18	C			

審査請求 未請求 請求項の数 1 O L (全 13 頁)

(21) 出願番号 特願平5-306398

(22) 出願日 平成5年(1993)12月7日

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 糸桐 健三

東京都品川区北品川6丁目7番35号 ソニー株式会社内

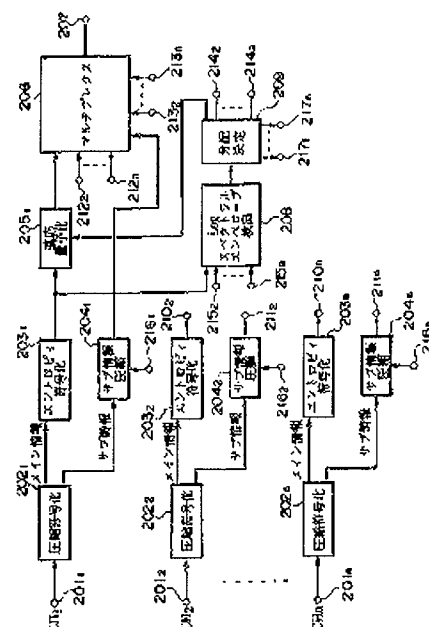
(74) 代理人 弁理士 小池 晃 (外2名)

(54) 【発明の名称】 多周符号化装置

(57) 【要約】

【構成】 各チャネルCH1～CHnのデジタルオーディオ信号を圧縮符号化する圧縮符号化回路202、～202n。及びエントロピ符号化回路203、～203n。と、メイン情報(圧縮符号化されたデジタルオーディオ信号)から各チャネル毎のエネルギーを検出してこの検出結果に応じた比率で各チャネルのビット配分を決定するLogスペクトラルエンベロープ検出回路208及び配分決定回路209と、この配分決定回路209からのビット配分情報に基づいて、適応量子化回路205、～205n。においてメイン情報を適応的に量子化すると共に、サブ情報(語長情報やスケールファクタの情報)を圧縮する。

【効果】 チャネル間ビット配分と高圧縮とを実現する。



(2)

特開平7-160292

1

【特許請求の範囲】

【請求項1】 複数チャンネルのデジタルオーディオ信号を多層符号化する多層符号化装置において、各チャンネルのデジタルオーディオ信号が入力され、当該デジタルオーディオ信号を圧縮符号化する第1の符号化手段と、上記第1の符号化手段の出力符号化信号が供給され、チャンネル毎のエネルギーを検出し、この検出出力に略応じた比率で、各チャンネルの情報量を割り当てる符号化を行う第2の符号化手段とを有してなることを特徴とする多層符号化装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、例えば、映画フィルム映写システム、ビデオテープレコーダ、ビデオディスクプレーヤ等のステレオや、いわゆるマルチサラウンド音響システムにおいて用いられるマルチチャンネルのデジタルオーディオ信号を圧縮符号化する多層符号化装置に関するものである。

【0002】

【従来の技術】オーディオ或いは音声等の信号の高効率符号化の手法及び装置には種々あるが、例えば、時間領域のオーディオ信号等を単位時間毎にブロック化してこのブロック毎の時間軸の信号を周波数軸上の信号に変換（直交変換）して複数の周波数帯域に分割し、各帯域毎に符号化するブロック化周波数帯域分割方式、いわゆる変換符号化（トランスフォームコーディング）や、時間領域のオーディオ信号等を単位時間毎にブロック化しないで、複数の周波数帯域に分割して符号化する非ブロック化周波数帯域分割方式である帯域分割符号化（サブバンドコーディング：SBC）等を挙げることができる。また、上述の帯域分割符号化と変換符号化とを組み合わせた高効率符号化の手法及び装置も考えられており、この場合には、例えば、上記帯域分割符号化で帯域分割を行った後、該各帯域毎の信号を周波数領域の信号に直交変換し、この直交変換された各帯域毎に符号化が施される。

【0003】ここで、上述した帯域分割符号化の帯域分割用フィルタとしては、例えばQMF等のフィルタがあり、これは1976 R.E.Crochiere Digital coding of speech in subbands Bell Syst.Tech. J. Vol.55, No.8 1976に、述べられている。また、ICASSP 83, BOSTON Polyphase Quadrature Filters-A new subband coding technique Joseph H. Rothweiler には等バンド幅のフィルタ分割手法及び装置が述べられている。

【0004】また、上述した直交変換としては、例えば、入力オーディオ信号を所定単位時間（フレーム）でブロック化し、該ブロック毎に高速フーリエ変換（FFT）、コサイン変換（DCT）、モディファイドDCT変換（MDCT）などを行うことで時間軸を周波数軸に

2

変換するような直交変換がある。このMDCTについては、ICASSP 1987 Subband/Transform Coding Using Filter Bank Designs Based on Time Domain Aliasing Cancellation J.P.Princen A.B.Bradley Univ. of Surrey Royal Melbourne Inst.of Tech.に述べられている。

【0005】更に、周波数帯域分割された各周波数成分を量子化する場合の周波数分割幅としては、例えば人間の聴覚特性を考慮した帯域分割がある。すなわち、一般に臨界帯域（クリティカルバンド）と呼ばれている高域程帯域幅が広くなるような帯域幅で、オーディオ信号を複数（例えば25バンド）の帯域に分割することがある。また、この時の各帯域毎のデータを符号化する際には、各帯域毎に所定のビット配分或いは、各帯域毎に適応的なビット配分による符号化が行われる。例えば、上記MDCT処理されて得られた係数データを上記ビット配分によって符号化する際には、上記各ブロック毎のMDCT処理により得られる各帯域毎のMDCT係数データに対して、適応的な配分ビット数で符号化が行われることになる。

【0006】上記ビット配分手法及びそのための装置としては、次の2手法及び装置が知られている。IEEE Transactions of Acoustics, Speech, and Signal Processing, vol. ASSP-25, No. 4, August 1977 では、各帯域毎の信号の大きさをもとに、ビット配分を行っている。また、ICASSP 1980 The critical band coder—digital encoding of the perceptual requirements of the auditory system M.A. Krasner MIT では、聴覚マスキングを利用することで、各帯域毎に必要な信号対雑音比を得て適応的なビット配分を行う手法及び装置が述べられている。

【0007】ここで、例えば上述したようなサブバンドコーディング等を用いたオーディオ信号の高効率圧縮符号化方式においては、人間の聴覚上の特性を利用し、オーディオデータを約1/5に圧縮するような方式が既に実用化されている。なお、このオーディオデータを約1/5に圧縮する高効率符号化方式としては、例えばいわゆるATRAC（Adaptive Transform Acoustic Coding）と呼ばれる方式が存在する。

【0008】さらに、通常のオーディオ機器の場合のみならず、例えば映画フィルム映写システム、高品位テレビジョン、ビデオテープレコーダ、ビデオディスクプレーヤ等のステレオないしはマルチサラウンド音響システムにおいては、例えば4～8チャンネル等の複数チャンネルのオーディオ或いは音声信号を扱うようになりつつあり、この場合においても、ビットレートを削減する高効率符号化を行うことが望まれている。

【0009】特に、業務用においては、デジタルオーディオのマルチチャンネル化が進んでおり、例えば8チャンネルのデジタルオーディオ信号を扱う機器が浸透してきている。上記8チャンネルのデジタルオーディオ信号

(3)

特開平7-160292

3

を録る機器としては、例えば映画フィルム映写システム等がある。また、高品位テレビジョン、ビデオテープレコーダ、ビデオディスクプレーヤ等のステレオないしはマルチサ라운드音響システムにおいても、例えば4～8チャンネル等の複数チャンネルのオーディオ或いは音声信号を録るようになりつつある。

【0010】ここで、上記8チャンネルのデジタルオーディオ信号を録る映画フィルム映写システムにおいては、上記映画フィルムに対して、例えばレフトチャンネル、レフトセンタチャンネル、センタチャンネル、ライトセンタチャンネル、ライトチャンネル、サラウンドレフトチャンネル、サラウンドライトチャンネル、サブウーファチャンネルの8チャンネルのデジタルオーディオ信号を記録することが行われつつある。なお、上記映画フィルムに記録する上記8チャンネルの各チャンネルは、例えば当該映画フィルムの画像記録領域から再生された画像が映写機によって投影されるスクリーン側に配置されるレフトスピーカ、レフトセンタスピーカ、センタスピーカ、ライトセンタスピーカ、ライトスピーカ、サブウーファスピーカ、観客席を取り囲むように左側に配置されるサラウンドレフトスピーカ及び右側に配置されるサラウンドライトスピーカと対応するものである。

【0011】ただし、映画フィルムに上記8チャンネルのデジタルオーディオ信号を記録する場合において、映画フィルムには、例えばいわゆるCD（コンパクトディスク）などで用いているようなサンプリング周波数44.1kHzで16ビットの直線量子化されたオーディオデータを上記8チャンネル分も記録できる領域を確保することは困難であるため、上記8チャンネルのオーディオデータを圧縮して記録するようになされる。例えば、当該8チャンネルのデジタルオーディオデータを圧縮する圧縮方法としては、上述したような人間の聴覚の特性を考慮して最適なビット割り当てを行うことによって、例えばいわゆるCD（コンパクトディスク）などに記録されるようなサンプリング周波数44.1kHzで16ビットのデジタルオーディオデータを約1/5に圧縮しながらも、CD並の音質を達成する前記高効率符号化方式（いわゆるATRAC方式など）を適用するようにしている。

【0012】また、フィルムという媒体は、表面に傷などが発生しやすいため、デジタルデータをオリジナルのまま記録していたのでは、データ欠けが激しく実用にならない。このため、エラー訂正符号の能力が非常に重要になり、上記データ圧縮は、その訂正符号も含めて上記フィルム上の記録領域に記録可能な程度まで行う必要がある。

【0013】

【発明が解決しようとする課題】しかし、上述のようなマルチチャンネルのオーディオ信号を高効率符号化する方式では、各チャンネル毎に圧縮が行われるため、全体とし

4

てのビット配分量（バイト配分量）が必ずしも有効に使用されているとは言い難い。例えば、あるチャンネルには少ないビット配分でもよいが、他のチャンネルではより多くのビット配分を必要とする場合がある。このように、従来の高効率符号化方式では、各チャンネルにわたってビット配分量（バイト配分量）を眺めると無駄が多いと思われる。特に、各チャンネル毎にビット配分量（或いはバイト配分量）が固定されている場合には、上記のような無駄がさらに顕著になると考えられる。近年は、さらに圧縮率を高めることも望まれている。

【0014】そこで、本発明は、上述したようなことに鑑み、より圧縮率を高めることができると共に、チャンネル間のビット配分をも可能とする多層符号化装置を提供することを目的としている。

【0015】

【課題を解決するための手段】本発明は、上述の目的を達成するために提案されたものであり、本発明の多層符号化装置は、複数チャンネルのデジタルオーディオ信号を多層符号化するものであり、各チャンネルのデジタルオーディオ信号が入力され、当該デジタルオーディオ信号を圧縮符号化する第1の符号化手段と、上記第1の符号化手段の出力符号化信号が供給され、チャンネル毎のエネルギーを検出し、この検出出力に略応じた比率で、各チャンネルの情報量を割り当てる符号化を行う第2の符号化手段とを有してなることを特徴とするものである。

【0016】

【作用】本発明によれば、第1の符号化手段で各チャンネルのデジタルオーディオ信号を圧縮符号化した後、この第1の符号化手段の出力符号化信号に対して、さらに第2の符号化手段によってチャンネル毎のエネルギーに略応じた比率で、各チャンネルの情報量を割り当てる符号化を行うことでチャンネル間のビット配分と高圧縮率を実現している。

【0017】

【実施例】以下、本発明の実施例について図面を参照しながら説明する。

【0018】図1には、本発明実施例の多層符号化装置の要部の構成を示す。

【0019】本発明実施例の多層符号化装置は、この図1に示すように、複数チャンネルCH1～CHnのデジタルオーディオ信号を多層符号化するものであり、各チャンネルCH1～CHnのデジタルオーディオ信号が入力され、当該デジタルオーディオ信号を圧縮符号化する第1の符号化手段としての圧縮符号化回路202、～202、及びエントロピ符号化回路203、～203、と、上記第1の符号化手段のエントロピ符号化回路203、～203、からのメイン情報の符号化信号（圧縮符号化回路202、～202、からの圧縮符号化されたデジタルオーディオ信号）が供給されてチャンネルCH1～CHn毎のエネルギーを検出するLogスペクトラ

(4)

特開平7-160292

5

ルエンベロープ検出回路208と、このLogスペクトラルエンベロープ検出回路208からの検出出力に略応じた比率で各チャンネルに割り当てる情報量を決定する

(ビット配分を行う)配分決定回路209と、この配分決定回路209からのビット配分情報に基づいて、適応変調回路205、～205。(回路205、～205。の図示は省略)において上記エントロピー符号化回路203、～203。からのメイン情報の符号化信号を適応的に変調すると共に、サブ情報圧縮回路204、～204。において上記圧縮符号化回路202、～202。からのデジタルオーディオ信号の圧縮符号化に関連するサブ情報(語長情報やスケールファクタの情報)を適応的に圧縮符号化する第2の符号化手段(Logスペクトラルエンベロープ検出回路208、配分決定回路209、適応変調回路203、～203。、サブ情報圧縮回路204、～204。、)とを有してなることを特徴とするものである。

【0020】なお、本発明の多層符号化装置によって圧縮符号化された各チャンネルのデジタルオーディオ信号は、伝送路を介して伝送されたり、記録媒体に記録されたりする。当該記録媒体として、例えば映画フィルムへの記録や、光ディスク、光磁気ディスク、相変化型光ディスク、磁気ディスク等のディスク状記録媒体、磁気テープ等のテープ状記録媒体への記録、半導体メモリ、ICカードなどへ記録される。

【0021】ここで、上記記録媒体として映画フィルムを用い、この映画フィルムへの記録を行う場合の上記各チャンネルすは1～C H 8は、例えば図2に示すようにスピーカが配置されるデジタルサラウンドシステムに対応することになる。各スピーカに対応するチャンネルは、センタ(C)チャンネル、サブウーファ(SW)チャンネル、レフト(L)チャンネル、レフトセンタ(CL)チャンネル、ライト(R)チャンネル、ライトセンタ(CR)チャンネル、レフトサラウンド(LB)チャンネル、ライトサラウンド(RB)チャンネルの8つである。

【0022】すなわちこの図2において、上記スピーカ配置に対応する各チャンネルは、例えば当該映画フィルムの画像記録領域から再生された画像が映写機(プロジェクタ100)によって投影されるスクリーン101側に配置されたレフトスピーカ106、レフトセンタスピーカ104、センタスピーカ102、ライトセンタスピーカ105、ライトスピーカ107、サラウンドレフトスピーカ108及び200、サラウンドライトスピーカ109及び201、サブウーファスピーカ103と対応するものである。

【0023】上記センタスピーカ102は、スクリーン101側の中央に配置され、センタチャンネルのオーディオデータによる再生音を出力するもので例えば俳優のせりふ等の最も重要な再生音を出力する。上記サブウーファスピーカ103は、サブウーファチャンネルのオーディオ

5

データによる再生音を出力するもので、例えば爆発音などの低域の音というよりは振動として感じられる音を効果的に出力するものであり、爆発シーンなどに効果的に使用されることが多いものである。上記レフトスピーカ106及びライトスピーカ107は、上記スクリーン101の左右に配置され、レフトチャンネルのオーディオデータによる再生音とライトチャンネルのオーディオデータによる再生音を出力するもので、ステレオ音響効果を発揮する。上記レフトセンタスピーカ104とライトセンタスピーカ105は、上記センタスピーカ102と上記レフトスピーカ106及びライトスピーカ107との間に配置され、レフトセンタチャンネルのオーディオデータによる再生音とライトセンタチャンネルのオーディオデータによる再生音を出力するもので、それぞれ上記レフトスピーカ106及びライトスピーカ107の補助的な役割を果たす。特にスクリーン101が大きく収容人数の多い映画館等では、座席の位置によって音像の定位が不安定になりやすいが、上記レフトセンタスピーカ104とライトセンタスピーカ107を付加することにより、音像のよりリアルな定位を作り出すのに効果を発揮する。さらに、上記サラウンドレフトスピーカ108とサラウンドライトスピーカ109は、観客席を取り囲むように配置され、サラウンドレフトチャンネルのオーディオデータによる再生音とサラウンドライトチャンネルのオーディオデータによる再生音を出力するもので、聴覚音や拍手、歓声に包まれた印象を与える効果がある。これにより、より立体的な音像を作り出すことができる。

【0024】図1に戻って、本実施例の多層符号化装置について説明する。この図1において、各入力端子201、～201。には、それぞれ対応する各チャンネルのデジタルオーディオ信号が供給される。

【0025】各入力端子201、～201。に供給された各チャンネルのデジタルオーディオ信号は、対応する各圧縮符号化回路202、～202。に送られて、ここで各チャンネル毎にそれぞれ圧縮符号化される。なお、各圧縮符号化回路202、～202。の詳細な構成については後述する。

【0026】各圧縮符号化回路202、～202。で圧縮符号化されたデジタルオーディオ信号は、メイン情報としてエントロピー符号化回路203、～203。に送られ、ここで、それぞれが可変長符号化される。また、上記圧縮符号化回路202、～202。でのデジタルオーディオ信号の圧縮符号化に関連する後述する語長情報やスケールファクタの情報は、サブ情報としてサブ情報圧縮回路204、～204。に送られ、ここで各チャンネル毎のビット配分情報に基づいて圧縮される。

【0027】上記エントロピー符号化回路203、～203。からのエントロピー符号化された各メイン情報は、適応変調回路205、～205。に送られると共に、Logスペクトラルエンベロープ検出回路208に

(5)

特開平7-160292

7

も送られる。なお、図1では、チャンネルCH2～CHnの各エントロピ符号化回路203<sub>2</sub>～203<sub>n</sub>からのエントロピ符号化されたメイン情報は、それぞれ対応する端子210<sub>2</sub>～210<sub>n</sub>から、図示を省略している各チャンネルCH2～CHnに対応する適応量子化回路205<sub>2</sub>～205<sub>n</sub>に送られると共に、端子215<sub>2</sub>～215<sub>n</sub>を介して上記Logスペクトラルエンベロープ検出回路208に送られるようになっている。

【0028】上記エントロピ符号化された各メイン情報が供給されるLogスペクトラルエンベロープ検出回路208では、これら各チャンネルのメイン情報からそれぞれエネルギーを検出して、各チャンネルのLogスペクトラルエンベロープを検出する。当該各チャンネルのLogスペクトラルエンベロープ情報は、上記配分決定回路209に送られる。

【0029】当該配分決定回路209では、上記Logスペクトラルエンベロープ検出回路208からの各チャンネルのLogスペクトラルエンベロープ情報に基づいて、各チャンネルのメイン情報に対するチャンネル間ビット配分情報と、各サブ情報に対するチャンネル間ビット配分情報を決定する。

【0030】上記配分決定回路209からの上記メイン情報に対するチャンネル間ビット配分情報は、適応量子化回路205<sub>2</sub>～205<sub>n</sub>に送られ、上記サブ情報に対するチャンネル間ビット配分情報は、上記サブ情報圧縮回路204<sub>2</sub>～204<sub>n</sub>に送られる。なお、図1の例では、チャンネルCH2～CHnのメイン情報に対するチャンネル間ビット配分情報は端子214<sub>2</sub>～214<sub>n</sub>を介して図示を省略している適応量子化回路205<sub>2</sub>～205<sub>n</sub>に送られ、チャンネルCH2～CHnのサブ情報に対するチャンネル間ビット配分情報は端子217<sub>2</sub>～217<sub>n</sub>及び端子216<sub>2</sub>～216<sub>n</sub>を介してサブ情報圧縮回路204<sub>2</sub>～204<sub>n</sub>に送られる。

【0031】上記適応量子化回路205<sub>2</sub>～205<sub>n</sub>では、上記メイン情報に対するチャンネル間ビット配分情報を用いて、上記各エントロピ符号化回路203<sub>2</sub>～203<sub>n</sub>からのエントロピ符号化されたデジタルオーディオ信号を適応的に量子化する。

【0032】また、上記サブ情報圧縮回路205<sub>2</sub>～205<sub>n</sub>では、上記サブ情報に対するチャンネル間ビット配分情報を用いて、上記圧縮符号化回路202<sub>2</sub>～202<sub>n</sub>からのサブ情報（語長情報、スケールファクタ情報）を圧縮符号化する。

【0033】上記適応量子化回路205<sub>2</sub>～205<sub>n</sub>とサブ情報圧縮回路205<sub>2</sub>～205<sub>n</sub>からの各出力は、マルチプレクサ206に送られる。なお、チャンネルCH2～CHnの図示を省略している適応量子化回路205<sub>2</sub>～205<sub>n</sub>からの各出力は、端子212<sub>2</sub>～212<sub>n</sub>を介して上記マルチプレクサ206に送られ、チャンネルCH2～CHnのビット情報圧縮回路205<sub>2</sub>～205<sub>n</sub>

8

5. からの各出力は、端子213<sub>2</sub>～213<sub>n</sub>を介して上記マルチプレクサ206に送られる。

【0034】当該マルチプレクサ206では、供給された各データをマルチプレクスして端子207から出力する。この出力データが符号化データとして例えば記録媒体等に記録される。

【0035】また、図1の本実施例装置においては、例えば図3に示すように、上記圧縮符号化回路202<sub>2</sub>～202<sub>n</sub>に対して供給する各チャンネルのデジタルオーディオ信号から、少なくとも2チャンネル毎（図3の例では全チャンネル毎）に対応してそれぞれ設けたハイパスフィルタ222によって高域成分のみを取り出した後、これら各高域成分を加算回路223によって加算し、端子225を介した当該高域成分の加算データに対して、所定の処理、例えば圧縮符号化、エントロピ符号化及び量子化の処理を施すようにすることもできる。このとき、上記高域成分を取り出す各チャンネルの入力デジタルオーディオ信号からは、それぞれローパスフィルタ221によって低域成分も取り出し、端子224を介して対応する圧縮符号化回路202<sub>2</sub>～202<sub>n</sub>に送るようにする。その後、上記高域及び低域成分を取り出す各チャンネルのうち、ある1つのチャンネル（図3の例ではチャンネルCH1）の上記マルチプレクサ206へ入力するメイン情報に対して、上記所定の処理を施した加算データを加算する処理（チャンネル間クロストーク処理）を行うようにする。

【0036】ここで、上記各チャンネルの高域成分を加算して、ある1つのチャンネルの低域成分に加算することを行うのは、以下の理由による。

【0037】すなわち、人間の耳は高域の成分に対する定位感が少なく、このため高域成分については例えば複数個のスピーカのうちいずれか1つのスピーカからでていても人間にはどのスピーカから出てきているのか聞き取り難いという性質がある。このため、複数チャンネルのオーディオ信号の各高域成分を、そのうちの例えば1チャンネルに対応するスピーカのみに送るようにしても、人間には各チャンネルのオーディオ信号の高域成分が当該1つのスピーカのみから出力されているとは感じられない。したがって、上述のように、上記各チャンネルの高域成分の音声を加算して、この加算データのある1つのチャンネルの低域成分に加算することで、各チャンネルの高域成分を1つのチャンネル分に圧縮できることになる。

【0038】次に、上記図1に示した圧縮符号化回路202<sub>2</sub>～202<sub>n</sub>の具体的な構成について説明する。

【0039】これら圧縮符号化回路202<sub>2</sub>～202<sub>n</sub>では、オーディオPCM信号等の入力デジタル信号を、帯域分割符号化（SBC）、適応変換符号化（ATC）及び適応ビット配分（APC-AB）の各技術を用いて圧縮符号化している。

【0040】以下、図4を用いて説明する。図4に示す

(6)

特開平7-160292

9

10

本実施例の圧縮符号化回路202、～202。では、入力デジタル信号をフィルタなどにより複数の周波数帯域に分割すると共に、各周波数帯域毎に直交変換を行って、得られた周波数軸のスペクトルデータを、後述する人間の聴覚特性を考慮したいわゆる臨界帯域幅（クリティカルバンド）毎に適応的にビット配分して符号化している。この時、高域では臨界帯域幅を更に分割した帯域を用いる。もちろんフィルタなどによる非ブロッキングの周波数分割幅は等分割幅としてもよい。さらに、本実施例においては、直交変換の前に入力信号に応じて適応的にブロックサイズ（ブロック長）を変化させると共に、クリティカルバンド単位もしくは高域では臨界帯域幅（クリティカルバンド）を更に細分化したブロックでフローティング処理を行っている。このクリティカルバンドとは、人間の聴覚特性を考慮して分割された周波数帯域であり、ある純音の周波数近傍の同じ強さの狭帯域バンドノイズによって当該純音がマスクされるときそのノイズの持つ帯域のことである。このクリティカルバンドは、高域ほど帯域幅が広くっており、上記0～22 kHzの全周波数帯域は例えば25のクリティカルバンドに分割されている。

【0041】すなわち、図4において、入力端子10には例えば0～22 kHzのオーディオPCM信号が供給されている。この入力信号は、例えばいわゆるQMF等の帯域分割フィルタ11により0～11 kHz帯域と11 kHz～22 kHz帯域とに分割され、0～11 kHz帯域の信号は同じくいわゆるQMF等の帯域分割フィルタ12により0～5.5 kHz帯域と5.5 kHz～11 kHz帯域とに分割される。帯域分割フィルタ11からの11 kHz～22 kHz帯域の信号は、直交変換回路の一例であるMDCT（Modified Discrete Cosine Transform）回路13に送られ、帯域分割フィルタ12からの5.5 kHz～11 kHz帯域の信号はMDCT回路14に送られ、帯域分割フィルタ12からの0～5.5 kHz帯域の信号はMDCT回路15に送られることにより、それぞれMDCT処理される。なお、各MDCT回路13、14、15では、各帯域毎に設けたブロック決定回路19、20、21により決定されたブロックサイズに基づいてMDCT処理がなされる。

【0042】ここで、上記ブロック決定回路19、20、21により決定される各MDCT回路13、14、15でのブロックサイズ的具体例を図5のA及びBに示す。なお、図5のAには直交変換ブロックサイズが長い場合（ロングモードにおける直交変換ブロックサイズ）を、図5のBには直交変換ブロックサイズが短い場合（ショートモードにおける直交変換ブロックサイズ）を示している。この図5の具体例においては、3つのフィルタ出力は、それぞれ2つの直交変換ブロックサイズを持つ。すなわち、低域側の0～5.5 kHz帯域の信号及び中域の5.5 kHz～11 kHz帯域の信号に対して

は、長いブロック長の場合（図5のA）は1ブロック内のサンプル数を128サンプルとし、短いブロックが選ばれた場合（図5のB）には1ブロック内のサンプル数を32サンプル毎のブロックとしている。これに対して高域側の11 kHz～22 kHz帯域の信号に対しては、長いブロック長の場合（図5のA）は1ブロック内のサンプル数を256サンプルとし、短いブロックが選ばれた場合（図5のB）には1ブロック内のサンプル数を32サンプル毎のブロックとしている。このようにして短いブロックが選ばれた場合には各帯域の直交変換ブロックのサンプル数を同じとして高域程時間分解能を上げ、なおかつブロック化に使用するウィンドウの傾斜を減らしている。なお、上記ブロック決定回路19、20、21で決定されたブロックサイズを示す情報は、後述の適応ビット配分符号化回路16、17、18に送られると共に、出力端子23、25、27から出力される。

【0043】再び図5において、各MDCT回路13、14、15にてMDCT処理されて得られた周波数領域のスペクトルデータあるいはMDCT係数データは、いわゆる臨界帯域（クリティカルバンド）または高域では更にクリティカルバンドを分割した帯域毎にまとめられて適応ビット配分符号化回路16、17、18に送られている。

【0044】適応ビット配分符号化回路16、17、18では、上記ブロックサイズの情報、及び臨界帯域（クリティカルバンド）または高域では更にクリティカルバンドを分割した帯域毎に割り当てられたビット数に応じて各スペクトルデータ（あるいはMDCT係数データ）を再量子化（正規化して量子化）するようにしている。適応ビット配分符号化回路16、17、18で符号化されたデータは、出力端子22、24、26を介して取り出される。また、当該適応ビット配分符号化回路16、17、18では、どのような信号の大きさに関する正規化がなされたかを示すスケールファクタ情報と、どのようなビット長で量子化がされたかを示すビット長情報（語長情報）も求めており、これらも同時に出力端子22、24、26から出力される。

【0045】また、図5における各MDCT回路13、14、15の出力からは、上記臨界帯域（クリティカルバンド）または高域では更にクリティカルバンドを分割した帯域毎のエネルギーが、例えば当該バンド内での各振幅値の2乗平均の平方根を計算すること等により求められる。もちろん、上記スケールファクタそのものを以後のビット配分の為に用いるようにしてもよい。この場合には新たなエネルギー計算の演算が不要となるため、ハード規模の節約となる。また、各バンド毎のエネルギーの代わりに、振幅値のピーク値、平均値等を用いることも可能である。

【0046】ここで、上記適応ビット配分回路16、17、18のより具体的な構成を図6で説明する。この図

11

6に示す適応ビット配分回路では、MDC T係数の大きさが各ブロック毎に求められ、そのMDC T係数が入力端子801に供給される。当該入力端子801に供給されたMDC T係数は、帯域毎のエネルギー算出回路803に与えられる。帯域毎のエネルギー算出回路803では、クリティカルバンド又は高域においてはクリティカルバンドを更に再分割したそれぞれの帯域に関する信号エネルギーを算出する。帯域毎のエネルギー算出回路803で算出されたそれぞれの帯域に関するエネルギーは、エネルギー依存ビット配分回路804に供給される。

【0047】エネルギー依存ビット配分回路804では、使用可能総ビット発生回路802からの使用可能総ビット、本実施例では128Kbpsの内のある割合（本実施例では100Kbps）を用いて白色の量子化雑音を作り出すようなビット配分を行う。このとき、入力信号のトーンリティが高いほど、すなわち入力信号のスペクトルの凸凹が大きいほど、このビット量が上記128Kbpsに占める割合が増加する。なお、入力信号のスペクトルの凸凹を検出するには、隣接するブロックのブロックフローティング係数の差の絶対値の和を指標として使う。そして、求められた使用可能なビット量につき、各帯域のエネルギーの対数値に比例したビット配分を行う。

【0048】聴覚許容雑音レベルに依存したビット配分算出回路805は、先ず上記クリティカルバンド毎に分割されたスペクトルデータに基づき、いわゆるマスキング効果等を考慮した各クリティカルバンド毎の許容ノイズ量を求め、次に聴覚許容雑音スペクトルを与えるように使用可能総ビットからエネルギー依存ビットを引いたビット分が配分される。このようにして求められたエネルギー依存ビットと聴覚許容雑音レベルに依存したビットは加算されて、図4の適応ビット配分符号化回路16、17、18により各クリティカルバンド毎もしくは高域においてはクリティカルバンドを更に複数帯域に分割した帯域に割り当てられたビット数に応じて各スペクトルデータ（あるいはMDC T係数データ）を再量子化するようにしている。このようにして符号化されたデータは、図4の出力端子22、24、26を介して取り出される。

【0049】さらに詳しく上記聴覚許容雑音スペクトル依存のビット配分回路805中の聴覚許容雑音スペクトル算出回路について説明すると、MDC T回路13、14、15で得られたMDC T係数が上記許容雑音算出回路に与えられる。

【0050】図7は上記許容雑音算出回路をまとめて説明した一具体例の概略構成を示すブロック回路図である。この図7において、入力端子521には、MDC T回路13、14、15からの周波数領域のスペクトルデータが供給されている。

【0051】この周波数領域の入力データは、帯域毎の

(7)

特開平7-160292

12

エネルギー算出回路522に送られて、上記クリティカルバンド（臨界帯域）毎のエネルギーが、例えば当該バンド内での各振幅値2乗の総和を計算すること等により求められる。この各バンド毎のエネルギーの代わりに、振幅値のピーク値、平均値等が用いられることもある。このエネルギー算出回路522からの出力として、例えば各バンドの総和値のスペクトルは、一般にパースペクトルと称されている。図8はこのような各クリティカルバンド毎のパースペクトルSBを示している。ただし、この図8では、図示を簡略化するため、上記クリティカルバンドのバンド数を12バンド（B1～B12）で表現している。

【0052】ここで、上記パースペクトルSBのいわゆるマスキングに於ける影響を考慮するために、該パースペクトルSBに所定の重み付け関数を掛けて加算するような畳込み（コンボリューション）処理を施す。このため、上記帯域毎のエネルギー算出回路522の出力すなわち該パースペクトルSBの各値は、畳込みフィルタ回路523に送られる。該畳込みフィルタ回路523は、例えば、入力データを順次遅延させる複数の遅延素子と、これら遅延素子からの出力にフィルタ係数（重み付け関数）を乗算する複数の乗算器（例えば各バンドに対応する25個の乗算器）と、各乗算器出力の総和をとる総和加算器とから構成されるものである。なお、上記マスキングとは、人間の聴覚上の特性により、ある信号によって他の信号がマスクされて聞こえなくなる現象をいうものであり、このマスキング効果には、時間領域のオーディオ信号による時間軸マスキング効果と、周波数領域の信号による同時刻マスキング効果とがある。これらのマスキング効果により、マスキングされる部分にノイズがあったとしても、このノイズは聞こえないことになる。このため、実際のオーディオ信号では、このマスキングされる範囲内のノイズは許容可能なノイズとされる。

【0053】ここで、上記畳込みフィルタ回路523の各乗算器の乗算係数（フィルタ係数）の一具体例を示すと、任意のバンドに対応する乗算器Mの係数を1とすると、乗算器M-1で係数0.15を、乗算器M-2で係数0.0019を、乗算器M-3で係数0.0000086を、乗算器M+1で係数0.4を、乗算器M+2で係数0.06を、乗算器M+3で係数0.007を各遅延素子の出力に乗算することにより、上記パースペクトルSBの畳込み処理が行われる。ただし、Mは1～25の任意の整数である。

【0054】次に、上記畳込みフィルタ回路523の出力は引算器524に送られる。該引算器524は、上記畳込んだ領域での後述する許容可能なノイズレベルに対応するレベル $\alpha$ を求めるものである。なお、当該許容可能なノイズレベル（許容ノイズレベル）に対応するレベル $\alpha$ は、後述するように、逆コンボリューション処理を



(8)

特開平7-160292

13

行うことによって、クリティカルバンドの各バンド毎の許容ノイズレベルとなるようなレベルである。ここで、上記計算器524には、上記レベル $\alpha$ を求めるための許容関数（マスキングレベルを表現する関数）が供給される。この許容関数を増減させることで上記レベル $\alpha$ の制御を行っている。当該許容関数は、次に説明するような $(n-a)$ 関数発生回路525から供給されているものである。

【0055】すなわち、許容ノイズレベルに対応するレベル $\alpha$ は、クリティカルバンドのバンドの低域から順に与えられる番号を1とすると、次の式で求めることができる。

$$\alpha = S - (n - a)$$

この式において、 $n$ 、 $a$ は定数で $a > 0$ 、 $S$ は畳込み処理されたパースペクトルの強度であり、式中 $(n-a)$ が許容関数となる。例として $n = 38$ 、 $a = -0.5$ を用いることができる。

【0056】このようにして、上記レベル $\alpha$ が求められ、このデータは、割算器526に伝送される。当該割算器526では、上記畳込みされた領域での上記レベル $\alpha$ を逆コンボリューションするためのものである。したがって、この逆コンボリューション処理を行うことにより、上記レベル $\alpha$ からマスキングスレッシュホールドが得られるようになる。すなわち、このマスキングスレッシュホールドが許容ノイズスペクトルとなる。なお、上記逆コンボリューション処理は、複雑な演算を必要とするが、本実施例では簡略化した割算器526を用いて逆コンボリューションを行っている。

【0057】次に、上記マスキングスレッシュホールドは、合成回路527を介して減算器528に伝送される。ここで、当該減算器528には、上記帯域毎のエネルギー検出回路522からの出力、すなわち前述したパースペクトルSが、遅延回路529を介して供給されている。したがって、この減算器528で上記マスキングスレッシュホールドとパースペクトルSとの減算演算が行われることで、図3に示すように、上記パースペクトルSは、該マスキングスレッシュホールドMSのレベルで示すレベル以下がマスキングされることになる。なお、遅延回路529は上記合成回路527以前の各回路での遅延量を考慮してエネルギー検出回路522からのパースペクトルSを遅延させるために設けられている。

【0058】当該減算器528からの出力は、許容雑音補正回路530を介し、出力端子531を介して取り出され、例えば配分ビット数情報が予め記憶されたROM等（図示せず）に送られる。このROM等は、上記減算回路528から許容雑音補正回路530を介して得られた出力（上記各バンドのエネルギーと上記ノイズレベル設定手段の出力との差分のレベル）に応じ、各バンド毎の配分ビット数情報を出力する。

14

【0059】このようにしてエネルギー依存ビットと聴覚許容雑音レベルに依存したビットは加算されてその配分ビット数情報を用いて当該適応ビット配分符号化回路16、17、18では符号化が行われる。

【0060】すなわち要約すれば、適応ビット配分符号化回路16、17、18では、上記クリティカルバンドの各バンド帯域（クリティカルバンド）毎もしくは高域においてはクリティカルバンドを更に複数帯域に分割した帯域のエネルギーもしくはピーク値と上記ノイズレベル設定手段の出力との差分のレベルに応じて配分されたビット数で上記各バンド毎のスペクトルデータを量子化することになる。

【0061】ところで、上述した合成回路527での合成の際には、最小可聴カーブ発生回路532から供給される図9に示すような人間の聴覚特性であるいわゆる最小可聴カーブRCを示すデータと、上記マスキングスレッシュホールドMSとを合成することができる。この最小可聴カーブにおいて、雑音絶対レベルがこの最小可聴カーブ以下ならば該雑音は聞こえないことになる。この最小可聴カーブは、コーディングが同じであっても例えば再生時の再生ボリュームの違いで異なるものとなが、現実的なデジタルシステムでは、例えば16ビットダイナミックレンジへの音楽のはいり方にはさほど違いがないので、例えば4kHz付近の最も耳に聞こえやすい周波数帯域の量子化雑音が聞こえないとすれば、他の周波数帯域ではこの最小可聴カーブのレベル以下の量子化雑音は聞こえないと考えられる。したがって、このように例えばシステムの持つダイナミックレンジの4kHz付近の雑音が聞こえない使い方をすると仮定し、この最小可聴カーブRCとマスキングスレッシュホールドMSとを共に合成することで許容ノイズレベルを得るようになると、この場合の許容ノイズレベルは、図9中の斜線で示す部分までとすることができるようになる。なお、本実施例では、上記最小可聴カーブの4kHzのレベルを、例えば20ビット相当の最低レベルに合わせている。また、この図9は、信号スペクトルSSも同時に示している。

【0062】また、上記許容雑音補正回路530では、補正情報出力回路533から送られてくる例えば等ラウドネスカーブの情報に基づいて、上記減算器528からの出力における許容雑音レベルを補正している。ここで、等ラウドネスカーブとは、人間の聴覚特性に関する特性曲線であり、例えば1kHzの純音と同じ大きさに聞こえる各周波数での音の音圧を求めて曲線で結んだもので、ラウドネスの等感度曲線とも呼ばれる。またこの等ラウドネス曲線は、図9に示した最小可聴カーブRCと略同じ曲線を描くものである。この等ラウドネス曲線においては、例えば4kHz付近では1kHzのところに音圧が8〜10dB下がっても1kHzと同じ大きさに聞こえ、逆に、50kHz付近では1kHzの音圧

(9)

特開平7-160292

15

よりも約15dB高くないと同じ大きさに聞こえない。このため、上記最小可聴カーブのレベルを越えた雑音（許容ノイズレベル）は、該等ラウドネス曲線に応じたカーブで与えられる周波数特性を持つようにするのが良いことがわかる。このようなことから、上記等ラウドネス曲線を考慮して上記許容ノイズレベルを修正することは、人間の聴覚特性に適合していることがわかる。

【0063】以上述べた聴覚許容雑音レベルに依存したスペクトル形状を使用可能総ビット128Kbpsの内のある割合を用いるビット配分で行く。この割合は入力信号のトナリティが高くなるほど減少する。

【0064】次にビット量分割手法について説明する。図6に戻って、MDC T回路出力が供給される入力端子801からの信号は、スペクトルの滑らかさ算出回路808にも与えられ、ここでスペクトルの滑らかさが算出される。本実施例では、信号スペクトルの絶対値の隣接値間の差の絶対値の和を、信号スペクトルの絶対値の和で割った値を、上記スペクトルの滑らかさとして算出している。

【0065】上記スペクトルの滑らかさ算出回路808の出力は、ビット分割率決定回路809に与えられ、ここでエネルギー依存のビット配分と、聴覚許容雑音スペクトルによるビット配分間のビット分割率とが決定される。ビット分割率はスペクトルの滑らかさ算出回路808の出力値が大きいほど、スペクトルの滑らかさが無いと考えて、エネルギー依存のビット配分よりも、聴覚許容雑音スペクトルによるビット配分に重点をおいたビット配分を行う。ビット分割率決定回路809は、それぞれエネルギー依存のビット配分及び聴覚許容雑音スペクトルによるビット配分の大きさをコントロールするマルチプライヤ811及び812に対してコントロール出力を送る。ここで、仮にスペクトルが滑らかであり、エネルギー依存のビット配分に重きを置くように、マルチプライヤ811へのビット分割率決定回路809の出力が0.8の値を取ったとき、マルチプライヤ812へのビット分割率決定回路809の出力は1-0.8=0.2とする。これら2つのマルチプライヤの出力はアダー806で足し合わされて最終的なビット配分情報となって、出力端子807から出力される。

【0066】次に、図10には、上記図4の圧縮符号化回路に対応する伸張復号化回路の構成を示す。

【0067】この図10において、各帯域の量子化されたMDC T係数は、入力端子122、124、126に与えられ、使用されたブロックサイズ情報は、入力端子123、125、127に与えられる。復号化回路116、117、118では適応ビット配分情報を用いてビット割当を解除する。

【0068】次に、IMDC T回路113、114、115では周波数領域の信号が時間領域の信号に変換される。これらの部分帯域の時間領域信号は、IQMF回路

16

112、111により、全体域信号に復号化される。その後、IQMF回路111、112で帯域合成がなされ、出力端子110から出力される。

【0069】

【発明の効果】上述したように、本発明の多層符号化装置においては、第1の符号化手段で各チャネルのデジタルオーディオ信号を圧縮符号化した後、この第1の符号化手段の出力符号化信号に対して、さらに第2の符号化手段によってチャネル毎のエネルギーに略応じた比率で、各チャネルの情報量を割り当てる符号化を行うことにより、チャネル間のビット配分と高圧縮率を可能としている。

【図面の簡単な説明】

【図1】本発明実施例の多層符号化装置の概略構成を示すブロック回路図である。

【図2】8チャネルデジタルサラウンドシステムにおけるスピーカの配置を説明するための図である。

【図3】チャネル間クロストークを行う主要構成要素の概略構成を示すブロック回路図である。

【図4】本実施例の圧縮符号化回路の構成例を示すブロック回路図である。

【図5】圧縮符号化回路での信号の周波数及び時間分割を示す図である。

【図6】情報信号の大きさ及び聴覚許容雑音スペクトルの2者を用いたビット配分手法を実現する構成を示すブロック回路図である。

【図7】許容雑音レベルを求める構成を示すブロック回路図である。

【図8】各帯域の信号レベルによるマスキングスレショールドの例を示す図である。

【図9】情報スペクトル、マスキングスレショールド、最小可聴限を示す図である。

【図10】本実施例の圧縮符号化回路に対応する伸張復号化回路の構成例を示すブロック回路図である。

【符号の説明】

202、～202、・・・圧縮符号化回路  
203、～203、・・・エントロピー符号化回路  
204、～204、・・・サブ情報圧縮回路  
205、～205、・・・適応量子化回路  
206、・・・マルチプレクサ  
208、・・・logスペクトラルエンベロープ検出回路  
209、・・・分配決定回路  
221、・・・ローパスフィルタ  
222、・・・ハイパスフィルタ  
223、・・・加算回路  
11、12、・・・帯域分割フィルタ  
13、14、15、・・・MDC T回路  
16、17、18、・・・適応ビット配分符号化回路  
19、20、21、・・・ブロックサイズ決定回路  
22、24、26、・・・符号化出力端子

(10)

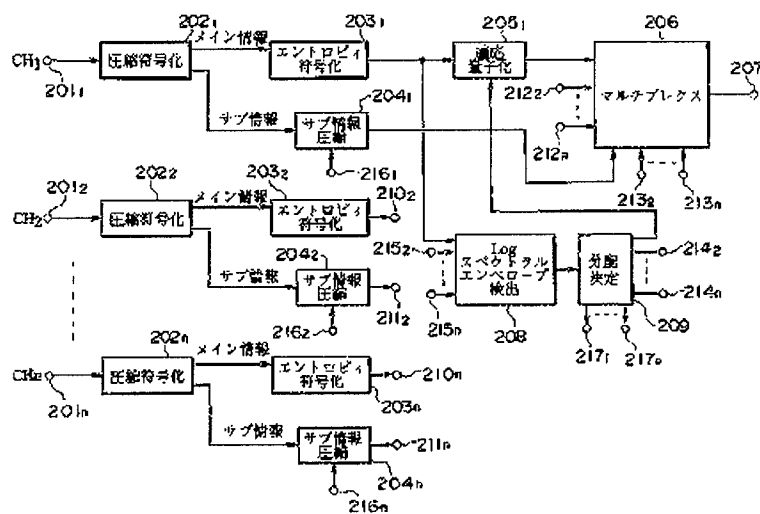
特開平 7-160292

17

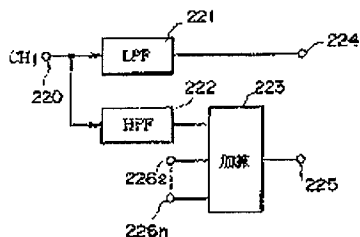
18

23, 25, 27・・・ブロックサイズ情報出力端子	* 528・・・減算器
116, 117, 118・・・適応ビット配分復号化回路	530・・・許容雑音補正回路
113, 114, 115・・・IMDCT回路	532・・・最小可聴カーブ発生回路
112, 111・・・IQMF回路	533・・・補正情報出力回路
520・・・許容雑音算出回路	801・・・MDC T回路出力端子
521・・・許容雑音算出回路入力端子	802・・・使用可能総ビット発生回路
522・・・帯域毎のエネルギー検出回路	803・・・帯域毎のエネルギー算出回路
523・・・畳込みフィルタ回路	804・・・エネルギー依存のビット配分回路
524・・・引算器	805・・・聴覚許容雑音レベル依存のビット配分回路
525・・・ $n-a_1$ 関数発生回路	806・・・アダー
526・・・割算器	807・・・各帯域のビット割当量出力端子
527・・・合成回路	808・・・スペクトルの滑らかさ算出回路
	809・・・ビット分割率決定回路
	* 811, 812・・・マルチプライヤ

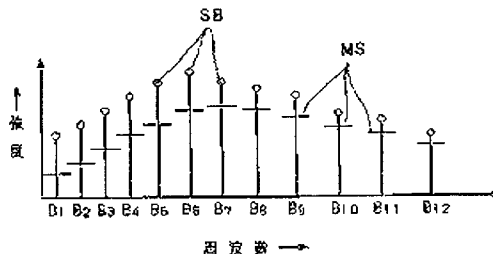
【図1】



【図3】



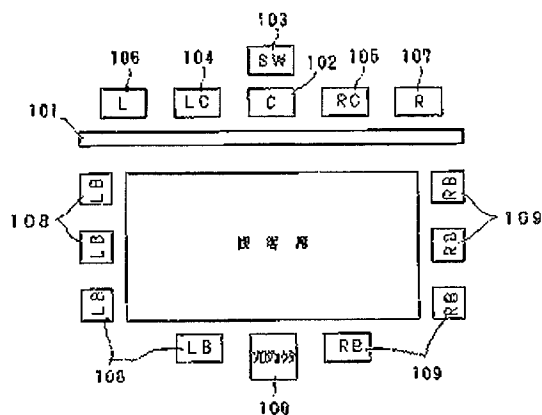
【図8】



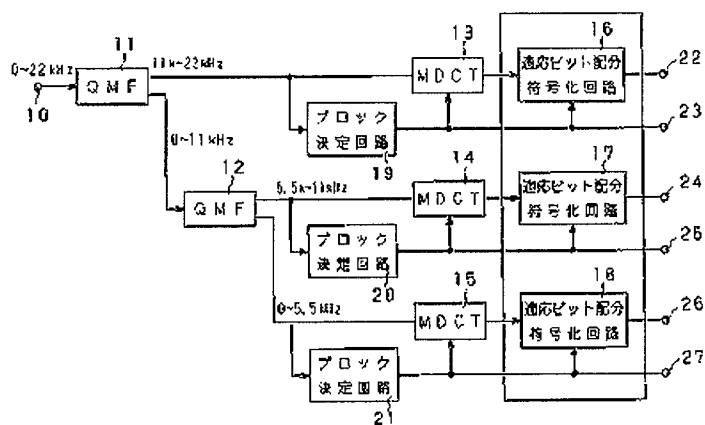
(11)

待開平? - 160292

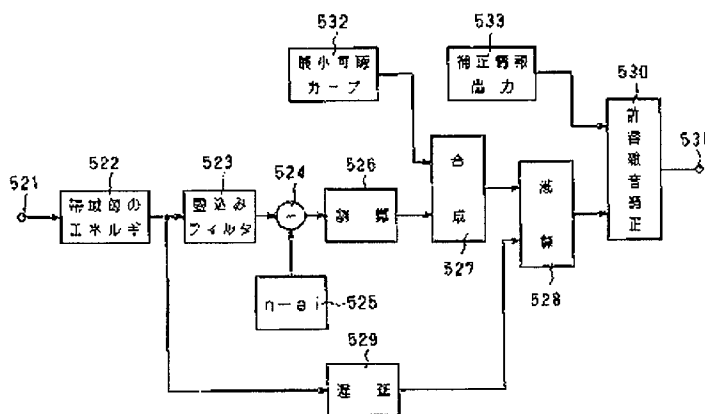
【圖 2】



【图4】



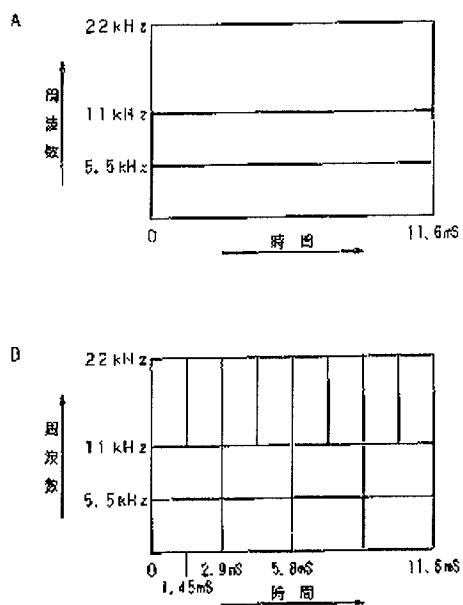
【図 7】



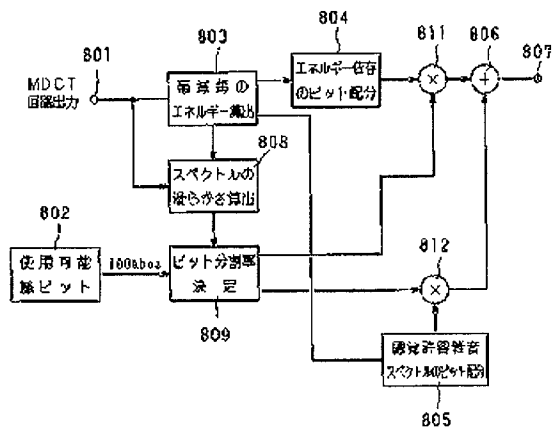
(12)

特開平7-160292

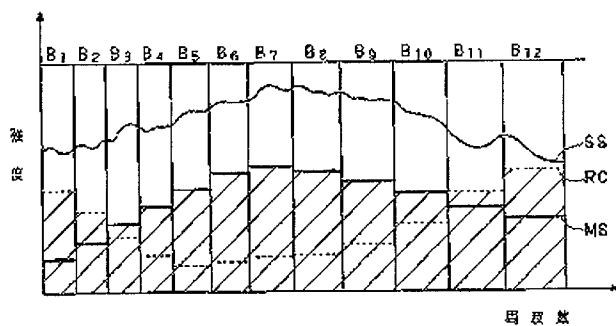
【図5】



【図6】



【図9】



(13)

特開平7-160292

【図10】

